Miniaturized all-solid-state 3D camera for real-time range imaging

Thierry Oggier*, Peter Seitz and Nicolas Blanc

CSEM SA, Swiss Center for Electronics and Microtechnology, Badenerstrasse 569

CH 8048 Zurich, Switzerland

ABSTRACT

The perception of the world in three-dimensions is natural for human beings. Technical 3D imaging systems, however, have suffered until today from high complexity and severe practical limitations to obtain 3D-information of the environment. This is overcome by a new type of optical 3D camera based on the time-of-flight (TOF) principle: Light from an LED or laser diode array is RF modulated at a few tens of MHz and illuminates a scene. The light is diffusely reflected back by the objects in the scene, and it is imaged with a conventional lens onto a custom solid-state image sensor. Each of its pixels is capable of synchronous demodulation of the incident modulated light, for the precise local determination of the parameters offset, amplitude and phase. The phase information is a direct measure of the local distance.

This principle has been employed in a miniaturized 3D camera (SwissRanger) for the acquisition of range images in video real-time. Without any mechanical scanning parts and with eye-safe emitting power, the camera delivers distance data, intensity information as well as an estimation of the distance accuracy for each of its 124x160 pixels.

Due to the use of a combination of CCD principles and CMOS circuitry in each pixel, a distance resolution is obtained that is close to the physical noise limitations given by the photon shot noise. Under optimum conditions a distance resolution of a few mm over a measurement range of several meters is obtained. A large number of applications are envisaged for which our TOF range camera provides a cost-effective and simple solution.

Keywords: *Optical 3D imaging, range finder, SwissRanger, distance measurement, lock-in pixel, time-of-flight, range camera.*

1. INTRODUCTION

We live in a three-dimensional world. Human beings are able to perceive their environment thanks to the "human stereo vision" system, the eyes. Several technical approaches are known to render imaging systems more intelligent by letting conventional cameras obtain information about the third dimension. Methods based on the triangulation principle were developed in the past e.g. stereo vision systems or systems based on structured projection. So far, the required computational cost underlying such a triangulation system and the inherent need of a minimum base distance of the two sensors (the projection and sensing unit) prevented their introduction into markets that are demanding low cost and small camera size. Another approach based on interferometry achieves very high distance resolutions, but due to its short measurement range, interferometers are only employed in a very restricted number of applications. Very promising results are achieved with time-of-flight (TOF) measuring devices. Resolutions in the sub-centimeter range and measurement ranges of several hundred meters have been reported. The drawback of these TOF-systems lies mainly in the high emitted power that is required and the necessity to use moving mechanical parts (e.g. scanners), leading to system costs that are unacceptable for many applications.

Until recently, no reliable and cost-effective 3D imaging cameras have been available on the market. The optical 3D camera described in this work was developed targeting a new, cost-effective imaging system allowing to capture three-dimensional imagery of the world in video real-time. A cost-effective and robust camera dictates the use of an all-solid-state, application-specific image sensor obviating the need for any moving parts. Extensive optoelectronic characterization results prove that the achieved distance accuracies approach the ultimate resolution given by the photon shot noise [1].

In the first section of this work, the principle of optical distance-measuring cameras based on the TOF method is explained. An electromagnetic wave, modulated at a few tens of MHz is emitted by the camera's illumination unit, illuminating the entire scene. The light, diffusely reflected by the different objects in the scene, is imaged onto the camera's custom sensor by a conventional lens. Within each pixel of this sensor, the signal is demodulated synchronously. The local phase shift - caused by the imaged object's distance - as well as the intensity and the amplitude of the incoming electro-magnetic wave are obtained at each pixel site.

In the following section, the working principle and the requirements of the sampling process in the so-called "lock-in pixel" [2] and the sensor are discussed. Extensive optoelectronic characterization results corroborate the theoretical predictions, proving that the SwissRanger camera yields distance data whose precision is close to the photon quantum noise limit.

In the fourth section different possible applications are described, in which the knowledge of the third dimension makes it possible to work with straightforward, simple algorithms. Some pictorial results of these algorithms illustrate the effectiveness of the chose approach based on range image data. The market potential of such cost-effective 3D-imaging devices in their different application fields is discussed and illustrated.

Finally an outlook for future improvements in the field of optical 3D cameras is given and their implications on the camera's performance is briefly discussed.

2. TOF MEASUREMENT PRINCIPLE

The TOF measurement principle is based on the finite speed of light $c \cong 3 \cdot 10^8\ ms^{-1}$ (in air). In order to reduce the required timing constraints on the system, instead of a pure pulse TOF measurement, a homodyne phase shift measurement is performed, as illustrated in Figure 1.

The principle is based on an illumination unit, which emits an intensity-modulated electromagnetic wave front. This emission signals can be described as

$$e(t) = e \cdot \left[1 + \sin\left(2\pi f \cdot t\right)\right] \qquad (1)$$

e: Emitted mean optical power
f: Modulation frequency

After reflection on the objects in the scene, the wave front reaches the sensor again. The power impinging on the sensor is reduced by different target characteristics (reflectivity, distance) and optical properties of the camera. The total signal attenuation can be summarized as a factor *k*. In addition to the modulated light, background light *BG* e.g. sun light or artificial light is sensed. Compared to the high frequency of the modulated signal, the background illumination can be considered as being constant. Therefore, the signal power impinging on a pixel can be described as

$$s(t) = BG + e \cdot k \cdot \left[1 + \sin\left(2\pi f \cdot t - \varphi\right)\right] \qquad (2)$$

BG: Background illumination power
k: Attenuation factor
φ: Phase delay arising from the object's distance

By sampling the incoming signal four times within a modulation period, the incident signal's modulation parameters can be completely determined, by making use of Equ. (3), (4), and (5). This calculation procedure is known as the four-bucket algorithm [3]. The four sampled and accumulated photo charge signals are denoted as A_0, A_1, A_2 and A_3.

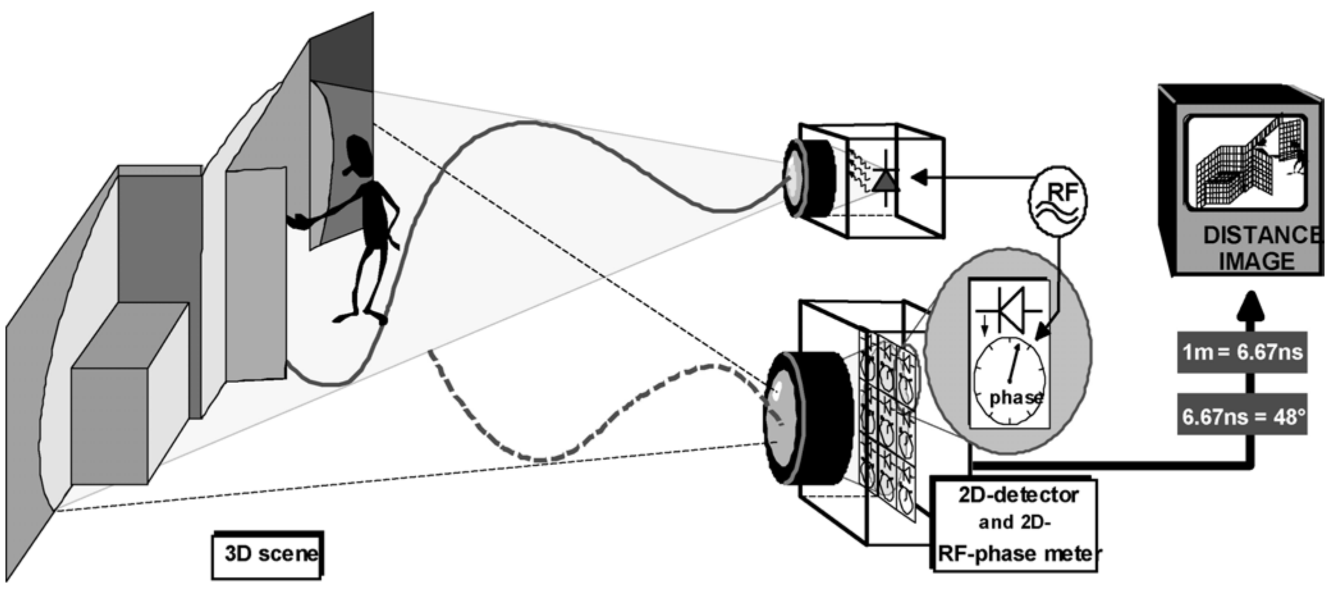

Figure 1: Measurement principle of the optical TOF 3D range camera. High-frequency modulated light is emitted by an LED array, it is reflected diffusely by the object, and it is demodulated synchronously in a lock-in image sensor.

$$\varphi = \operatorname{atan}\left(\frac{A_3 - A_1}{A_0 - A_2}\right) \quad (3)$$

$$B = \frac{A_0 + A_1 + A_2 + A_3}{4} \quad (4)$$

$$A = \frac{\sqrt{[A_3 - A_1]^2 + [A_0 - A_2]^2}}{2} \quad (5)$$

φ: Measured phase delay
B: Measured offset
A: Measured amplitude

An illustration of the sampling process is sketched in figure 2. The phase φ represents a direct measure of the acquired target distance, *B* corresponds to a conventional black and white intensity image and *A* is the amplitude of the incoming wave.

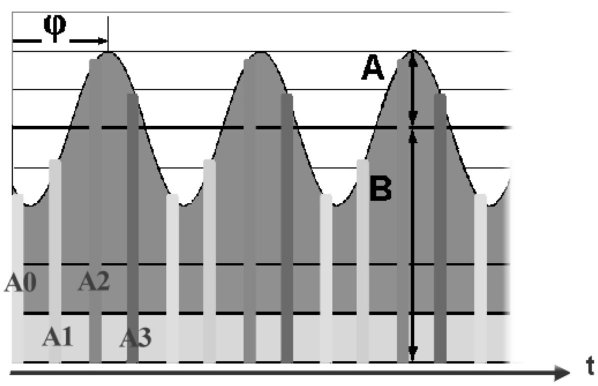

Figure 2: Illustration of the four samplings of the wave A0, A1, A2 and A3 and the reconstituted wave characteristics amplitude A, intensity B and phase φ.

The distance can be derived from the phase according to

$$L = \frac{L_0}{2\pi} \cdot \varphi \quad (6)$$

L_0: Non-ambiguity range

L_0 represents the non-ambiguity range of the phase measurement and corresponds to half the wavelength of the modulation frequency (7).

$$L_0 = \frac{c}{2f} \quad (7)$$

A more detailed discussion of homodyne phase measurement and its physical detection limit given by the photon shot noise is given in [4].

3. SWISSRANGER CAMERA DEMONSTRATOR

Figure 1 illustrates the different components required for the distance-measuring device. The camera consists of:

- Illumination unit
- Optical elements for imaging and filtering
- Custom sensor
- Sensor control electronics
- Data processing electronics
- Camera interface

Based on the described synchronous demodulation image sensor and the optical TOF range imaging principle, a miniaturized camera demonstrator (SwissRanger) was developed, shown in Fig. 3. Apart from the custom-made sensor, the SwissRanger camera only consists of commercially available components.

Its dimensions are 135 mm x 45 mm x 32 mm and the camera weighs less than 200 g, of which more than half is the weight of the metal case. The camera offers a lateral resolution of 160 x 124 pixels and during operation it consumes about 1.5 A at 12 V, depending on the particular measurement settings. Due to the use of a solid-state image sensor and the commercial availability of all other microelectronic components, the camera's measurement principle is well suited for the fabrication of cost-effective products.

The illumination unit emits an intensity-modulated light wave of a few 10 MHz into the entire field-of-view (FOV). The total emitted mean power amounts to up to 800 mW. LEDs emitting at a wavelength of 870 nm are employed. In general, laser diodes or LEDs can be used in the illumination unit. In any case, the maximal optical power limit is given by eye safety considerations. Implementing diffusive micro-optical elements in front of the emitting diodes eases on the one hand the eye safety limitations and on the other hand illuminates the FOV in a more controlled manner.

Figure 3: Photograph of the SwissRanger camera demonstrator.

The optical elements consist of the imaging optics in front of a bandpass filter. The imaging optics forms an image of the scene in the sensor plane, and the filter reduces possible background illumination in the scene. The imaging optics has to be adjusted carefully to the illumination unit and the sensor. Ideally, the illumination unit only emits light in the FOV determined by the optics and the sensor. The filter also needs to be designed according to the light source. LEDs require a broader bandpass filter than laser diodes. Therefore, for outdoor applications laser diodes allowing the implementation of a more narrow bandpass filter are preferred to LEDs.

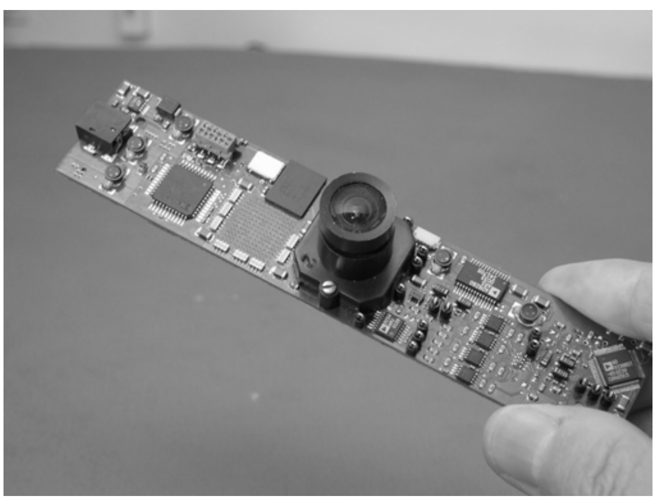

Figure 4: Electronics board of the SwissRanger camera demonstrator.

The electronics board provides the different signals to control the sensor. These signals are required for the readout, supply voltages or gate controls. The analog to digital conversion is performed on the printed circuit board, as well as the entire data processing, implementing the equations (3), (4) and (5). The camera settings allow the definition of a noise-level threshold to withhold inaccurate distance measurements from being made available outside the camera. The interface to the computer conforms to the widely used USB2.0 standard. Different parameters such as integration time, definition of region-of-interest (ROI) or spatial filtering can be programmed by the user through the USB2.0 interface.

4. 3D LOCK-IN IMAGE SENSOR

The application-specific image sensor has been designed, simulated and manufactured in a 0.8 μm CMOS/CCD technology from the silicon foundry ZMD in Dresden, Germany. The sensor contains 160 x 124 pixels. Each pixel can be addressed and read out individually, thus, any arbitrary ROI can be defined. The sensor is based on the so-called 2-tap pixel architecture, implying that each pixel contains two photocharge storage sites. This architecture represents a trade-off between speed (4-tap) and sensitivity (1-tap). A more detailed comparison about the different multi-tap pixel architectures is reported in [5].

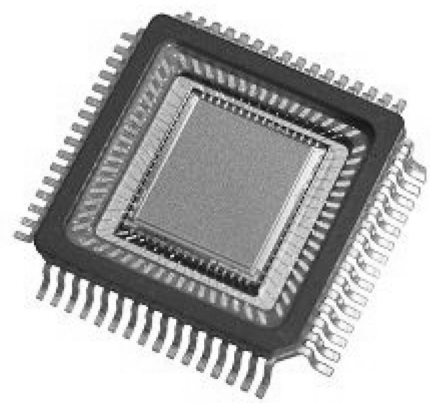

Figure 5: Picture of the SwissRanger sensor manufactured on a 0.8 μm CMOS/CCD technology.

The pixel has been created using ISE-TCAD simulation tools. The CCD-gates within each pixel allow fast photocharge separation, temporal sampling and accumulation. More details about CCD imaging have been published in [6] and [7]. Because of the high modulation frequencies and the requirements in TOF-systems regarding the shutter, dominating the demodulation characteristics, it is of highest importance to transfer and sum the photocharges with as little additional noise as possible. For this reason, we employ the CCD principle, realized with a buried channel CCD option in an otherwise standard CMOS process for virtually noise-free photocharge separation, sampling and accumulation. The same CMOS technology is used to address and read out the pixels.

This combination of CMOS and CCD technology does not only result in a pixel performance that is optimized for optical TOF range imaging, it also opens the way to future system-on-chip (SoC) solutions of cost-effective single-chip TOF range cameras.

Extensive optoelectronic characterization confirms that the SwissRanger camera demonstrator approaches the physical detection limit given by the photon shot noise. Under favorable measurement conditions (little background illumination and pixel signals close to saturation) a distance resolution of a few mm can be obtained for a measurement range of several meters [1].

5. APPLICATION EXAMPLE

Thanks to the advantageous properties of the SwissRanger camera (high distance resolution of TOF-measurements, high lateral resolution and speed) completely new solutions to difficult measurement problems can be envisaged. A large number of applications fields can be covered in a more cost-efficient way.

In this paragraph, we describe a few application examples of optical 3D range cameras based on the TOF principle. Typical pictures illustrate the feasibility of the 3D imaging approach for these problems.

Figure 6 shows three views of one 3D picture of a human face taken by a SwissRanger camera.

Figure 6: Shot of the SwissRanger camera on a human face.

The 3D picture was taken with an exposure time of less than one second. The black-and-white reflectivity information is projected onto the measured three-dimensional surface. This results implies that optical 3D imaging can provide valuable additional information to the reflectivity map of a human face, without any special requirements, in a short time and at low cost. It seems obvious that biometrical applications could make good use of this additional information for increased security.

Another application example is motion tracking. The knowledge of the third dimension allows implementing much faster and easier detection algorithms. Figure 7 shows the hand tracking of a person in a room. Three pictures of a video-rate sequence are presented.

Figure 7: 1st row: b/w coded distance map. 2nd row: b/w reflectance image. 3rd row: hand tracking.

In the first row of figure 7, the distance map is intensity coded. Pixels considered not accurate enough for the hand tracking are withheld by the camera at pixel level. These pixels are represented in the first and second row as black pixels. Pixels in the first image row represent the measured distance. Dark pixels correspond to closer distance whereas bright pixels originate from objects farther away from the camera. The second row shows the conventional b/w reflectance images of the sequence. The third row shows those pixels in white that are extracted by a simple hand-tracking algorithm.

Information about the third dimension makes it much easier to develop fast, reliable algorithms of low complexity, for which the presented hand tracking is a good example: Algorithms based on the distance maps save a lot of computational power compared to pure 2D algorithms.

Figure 8 shows a sequence of range images of a scene, acquired at video-rate. The distance on the different maps is b/w coded. Objects with darker pixels are closer to the camera; brighter objects are situated farther from the camera. For illustration, the used colour-coded distance scale is provided. The sequence shows a person entering a room. This person can easily be tracked and detected by processing the distance information. Any possible influences by lighting effects, such as shadows, do not influence the distance map, in strong contrast to the detection algorithms based on conventional 2D-sensors.

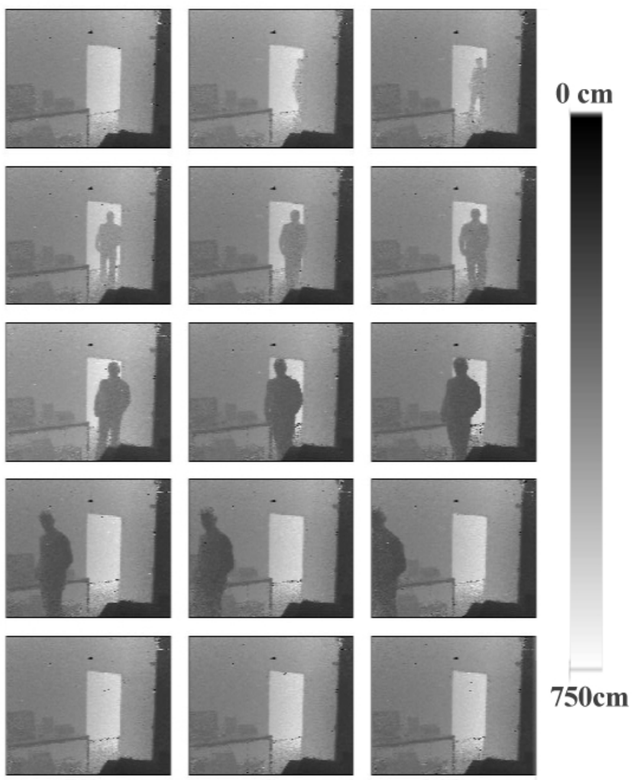

Figure 8: Sequence of distance maps acquired by the SR-2 camera for room observation.

6. OUTLOOK

The next generation of 3D-TOF cameras will be based on CMOS/CCD technologies with smaller feature sizes. This will increase the number of lock-in pixels and their density on one custom TOF image sensor. By using laser diodes and vertical cavity surface emitting lasers (VCSEL), outdoor applications are targeted, first at closer distances and then at distances of several tens of meters. New pixel architectures and technologies will allow higher modulation frequencies [8] and thus further improve the distance resolution. Finally, the use of microlenses is foreseen in order to increase the pixel sensitivity.

7. SUMMARY AND CONCLUSIONS

In this paper, an overview of the optical homodyne phase measurement technique used for distance measurements is given. The TOF-principle as implemented in the SwissRanger camera demonstrator is presented and the different required components are described.

The potential of the optical 3D TOF imaging principle is illustrated with three examples, face recognition, hand tracking and room observation, requiring only straight-forward and simple algorithms. Many more applications are foreseen in a wide range of applications, such as automotive, security, safety, door-and-gate control, robotics, autonomous vehicles, computer pointing devices, domotics, games, biometrics, toys, etc.

It is concluded that the described optical 3D imaging as implemented in the SwissRanger camera demonstrator represents an outstanding measurement concept for acquiring true images of our rich three-dimensional world, that can finally been acquired at high speed, high resolution, with miniature camera systems and at low cost.

ACKNOWLEDGMENT

The authors would like to thank all the members of the Image Sensing Group of CSEM SA for their continuing support in the various phases of this research and development project.

REFERENCES

[1] T. Oggier et al., "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger^{TM})", *Proceedings of the SPIE,* Vol. 5249 No. 65, 2003.

[2] T. Spirig, P. Seitz, O. Vietze, F. Heitger, „The Lock-in CCD – Two-Dimensional Detection of Light“, *IEEE journal of quantum electronics*, Vol.31, No. 9, 1995.

[3] K. Creath, "Phase-Measurement Interferometry Techniques", *Progress in Optics,* Vol. XXVI, E. Wolf (Ed.), Elsevier, 1988.

[4] R. Lange and P. Seitz, "Solid-State Time-of-Flight Range Camera", *IEEE J. Quantum Electronics*, Vol. 37 (3), 390-397, March 2001.

[5] W. Boyle and G. Smith, "Charge coupled semiconductor devices", *Bell Syst. Tech. Jour*., Vol. 49, pp. 587-593, 1970.

[6] R. Kaufmann et al., “A Time-of-Flight Line Sensor – Development and Application“, *Proceedings of the SPIE*, Vol. 5459 No. 24, 2004.

[7] A. Theuwissen, “Solid-State Imaging with Charge-Coupled Devices”, *Kluwer Academic Publishers*, 1995.

[8] B. Büttgen, T. Oggier, R. Kaufmann, P. Seitz, N. Blanc, “Demonstration of a novel drift field pixel structure for the demodulation of modulated light waves with application in three-dimensional image capture”, *Proceedings of the SPIE,* Vol. 5302 No. 2, 2004

---

* Contact: thierry.oggier@csem.ch